

Universal and Near-Universal Cycles of Set Partitions

Zach Higgins

Department of Mathematics
University of Florida

Elizabeth Kelley

Department of Mathematics
Harvey Mudd College

Bertilla Sieben

Department of Mathematics
Princeton University

Anant Godbole

Department of Mathematics and Statistics
East Tennessee State University

February 16, 2015

Abstract

We study universal cycles of the set $\mathcal{P}(n, k)$ of k -partitions of the set $[n] := \{1, 2, \dots, n\}$ and prove that the transition digraph associated with $\mathcal{P}(n, k)$ is Eulerian. But this does not imply that universal cycles (or ucycles) exist, since vertices represent equivalence classes of partitions! We use this result to prove, however, that ucycles of $\mathcal{P}(n, k)$ exist for all $n \geq 3$ when $k = 2$. We reprove that they exist for odd n when $k = n - 1$ and that they do not exist for even n when $k = n - 1$. An infinite family of (n, k) for which ucycles do not exist is shown to be those pairs for which $S(n - 2, k - 2)$ is odd ($3 \leq k < n - 1$). We also show that there exist universal cycles of partitions of $[n]$ into k subsets of distinct sizes when k is sufficiently smaller than n , and therefore that there exist universal packings of the partitions in $\mathcal{P}(n, k)$. An analogous result for coverings completes the investigation.

1 Introduction

Universal cycles are often loosely defined. In a recent seminar talk at Virginia Commonwealth University, Glenn Hurlbert offered the following description: “Broadly, universal cycles (ucycles) are special listings of combinatorial objects in which codes for the objects are written in an overlapping, cyclic manner.” By “special,” Hurlbert means “without repetitions”. As an example, the cyclic string 112233 encodes each of the six multisets of size 2 from the set $\{1, 2, 3\}$. Another well-quoted example, from [8], is the string

1356725 6823472 3578147 8245614 5712361 2467836 7134582 4681258,

where each block is obtained from the previous one by addition of 5 modulo 8. This string is an encoding of the fifty six 3-subsets of the set $[8] := \{1, 2, 3, 4, 5, 6, 7, 8\}$. A seminal paper in the area is that of Chung, Diaconis and Graham [2] who studied ucycles of

- subsets of size k of an n -element set (as in the above example);
- set partitions (the focus of this paper); and
- permutations (with a necessarily augmented ground set and the use of order isomorphism representations, e.g., the string 124324 encodes each of the six permutations of $[3] = \{1, 2, 3\}$ in an order isomorphic fashion, which is clearly not possible to do using the ground set $[3]$).

In [2] it is shown that for $n \geq 4$, there *does exist* a ucycle of all partitions $\mathcal{P}(n)$ of the set $[n]$ into an arbitrary number of parts; for example, we have the ucycle *abcbccccddcdeec* of $\mathcal{P}(4)$, where, for example, the substring *dcde* encodes the partition $13|2|4$. Note that the alphabet used was in this case of size 5, though an alphabet of (minimum) size 5 is shown to suffice to encode $\mathcal{P}(5)$ as

DDDDDCHHHCCDDCCCHCHCSHHSDSSDSSHSDDDCH
SSCHSHDHSCHSJCDC.

The above example reflects tongue-in-cheek humor, since there are 52 partitions of $[5]$ and the above ucycle has 13 cards of each suit – except that one spade has been replaced by a joker! The authors also ask how many partitions of $\mathcal{P}(n)$ using an alphabet of size $N \geq n$ exist. This question will

also be of deep relevance to us, as alluded to in the later section on Open Problems.

As noted in [6], however, not much seems to be known about ucycles of the partitions $\mathcal{P}(n, k)$ ($|\mathcal{P}(n, k)| = S(n, k)$) of $[n]$ into k parts. In [4], it was shown that for $k = n - 1$, ucycles exist if and only if n is odd. At the other end of the k -spectrum, the authors of [7] showed that for odd n , one could find a ucycle of partitions of $[n]$ into two parts, and that an “asymptotically good universal packing” could be found for $k = 3$, i.e., that there was a string of length $T(n, 3) < S(n, 3)$ with

$$\frac{T(n, 3)}{S(n, 3)} \rightarrow 1, \quad n \rightarrow \infty,$$

where each of the $T(n, 3)$ consecutive strings of length n represented a different partition of $[n]$ into 3 parts of distinct sizes. The authors of [7] also proved that ucycles of partitions of $[n]$ into 2 parts exist if $n = 4$ or 6, as evidenced by the respective explicit examples

$$aabbaba$$

and

$$abaaaabbaababbbabbbaaabbbbababa.$$

It is this work that we build on. In Section 2, we generalize the above result on asymptotically good universal packings (upackings) to the case of all fixed k as $n \rightarrow \infty$, as well as to ucoverings, which are also shown to be “asymptotically good.” Finally, Section 3 is devoted to showing that the transition digraph associated with $\mathcal{P}(n, k)$ is Eulerian. As noted in the work of [2], this does not necessarily imply that universal cycles exist, since the digraph vertices represent equivalence classes of partitions. We use our result to prove, however, that ucycles of $\mathcal{P}(n, k)$ exist for all $n \geq 3$ when $k = 2$ and for odd n when $k = n - 1$, the latter recovering the result in [4]. We also (re)prove that ucycles do not exist for n even when $k = n - 1$. Finally, we show that for even n , ucycles do not exist when $S(n - 2, k - 2)$ is odd ($3 \leq k < n - 1$). The last result shows, e.g., that ucycles of $\mathcal{P}(12, 6)$, or $\mathcal{P}(6, 3)$ do not exist. There are infinitely many such pairs of values of (n, k) . Moreover, the technique we exhibit in Section 3 has the potential to tease out many more results along these lines.

2 Universal Packings and Coverings of Partitions of $[n]$ into k parts

One of the main results in [9] was that one could create a ucycle of all surjections from $[n]$ to $[k]$ iff $n > k$. Since there are $k!$ surjections that yield the same set partition, we need to be more careful, and proceed by showing in Theorem 1 that for sufficiently large n , it is possible to ucycle partitions of $[n]$ into k parts of *distinct* sizes. We represent such partitions as surjections from $f : [n] \rightarrow [k]$; $n > k$, with $|f^{-1}(\{1\})| < |f^{-1}(\{2\})| < \dots < |f^{-1}(\{k\})|$. The fact that *asymptotically good upackings* exist is proved in Theorem 2, where we provide an alternative proof of the intuitively obvious fact that for any fixed k , the number of k -partitions of $[n]$ with non-distinct part sizes form a vanishing fraction of all partitions into k parts as $n \rightarrow \infty$.

Theorem 1. *For each fixed $k, n; n \geq \frac{(k+4)(k-1)}{2} + 1$, there exists a ucycle of all onto functions $f : [n] \rightarrow [k]$ such that the preimage cardinality function $|f^{-1}| : \{\{1\}, \{2\}, \dots, \{k\}\} \rightarrow [n]$ is strictly increasing.*

Proof. Following the standard process, we create a digraph D in which the vertices are sequences of length $n - 1$, of numbers in $\{1, \dots, k\}$, for which the addition of at least one number in $\{1, \dots, k\}$ at the beginning or the end creates a sequence of length n which, using the special canonical format we have adopted, represents a partition of $[n]$ into k parts of distinct sizes. For example, with $n = 10, k = 4$, 122333444 is a legal vertex, as is 123334444. On the other hand, 112233344 is not in the underlying graph. A vertex points towards another if its last $n - 2$ terms are the same as the first $n - 2$ terms of the second vertex. The edges of the digraph, labeled by vertex label concatenation, are thus sequences representing partitions of $[n]$ into k distinct parts.

The problem of finding a ucycle is reduced to the problem of finding an Eulerian circuit in this digraph. We know that Eulerian circuits exist if the graph is both strongly connected and balanced, i.e., for each vertex v , the in- and out-degrees of v are equal. It is easy to show that if a digraph is balanced and weakly connected, then it is also strongly connected, so all we need to show is that our digraph is balanced and weakly connected. This approach is used, e.g., in [1].

Showing that D is balanced is simple - if a number from 1 through k can be added to the beginning of the sequence at a vertex, then it can also

always be added to the end of the sequence since it is only the numbers of 1's, 2's, ... and k 's that actually matter in determining if an edge represents a partition into distinct parts. Therefore, the number of edges pointing away from a vertex will be the same as the number of edges pointing towards it. Note that the in- and out-degrees of any vertex (the common value is sometimes called the vertex degree), though equal, are quite different at different vertices. For example for $k = 3$, the vertex 123333333 has degree 1; the vertex 122333333 has degree 2; and the vertex 122233333 has degree 3. In general, one may write down a formula for $\deg(v)$ depending on the differences between the number of $i + 1$ s and the number of i s in v .

To show that the digraph is weakly connected, we need only show that it is possible to reach a designated target vertex from any other starting vertex. We will let this target vertex be the one consisting of two 2's, three 3's, ..., and $k - 1$ $k - 1$'s – leaving all of the remaining numbers as k 's, in that order. For example, for $n = 27$ and $k = 6$, the target vertex of length 26 will be 223334444555556666666666666666. In fact, we will show something stronger, namely that one can traverse from any edge to the edge $122333 \dots (k - 1) \dots (k - 1)k \dots k$ from whence we may reach the target vertex in a single step. Notice that such edges represent legal partitions in $\mathcal{P}(n, k)$ only if $n \geq n_0 := k(k + 1)/2$.

Key to our algorithm on how to reach one edge from another is the process of “switching” numbers. For example, we can write out all of the steps to go from the edge 33132323 to the edge 33122323 as follows $33132323 \rightarrow 31323233 \rightarrow 13232333 \rightarrow 32323331 \rightarrow 23233312 \rightarrow 32333122 \rightarrow 23331223 \rightarrow 33312232 \rightarrow 33122323$, or, we can skip all the steps of “rotating around” and just say that we “switched” the 3 to the right of the 1 into a 2. Note that “rotating” is always legal but switching in the above form might not always be, even in several steps. We need to have the “room to maneuver around” while maintaining edge-integrity.

We will show that the only requirement to reach the designated target edge from any other edge is to have the ability to “switch” any $j \in \{2, 3, \dots, k - 1\}$ into a k and vice versa (possibly through several steps), and that this is equivalent to having $n \geq n_1 := \frac{(k+4)(k-1)}{2} + 1 = n_0 + (k - 1)$, where the “extra” $(k - 1)$ digits give us the needed room to maneuver around.

We will let the partition size vector (PSV) of an edge be a k -tuple that expresses the number of 1's, 2's, etc. in order. For example, with $n = 18$; $k = 5$, the only possible PSV's are $(1, 2, 3, 4, 8)$; $(1, 2, 3, 5, 7)$; and $(1, 2, 4, 5, 6)$. In

this case, there is no way, e.g., to change a 2 with any other number. If, for example, a 2 is switched with a 5, this forces the numbers of 1's and 2's to be equal, and other issues might arise if a 2 is switched with a 3 or a 4. However, if we have $n = 19$ and $k = 5$, the PSV's are $(1, 2, 3, 4, 9)$, $(1, 2, 3, 5, 8)$, $(1, 2, 3, 6, 7)$, $(1, 2, 4, 5, 7)$, and $(1, 3, 4, 5, 6)$, which will be seen to imply that no matter what our starting position, we have enough "spaces" so that we can switch a 2, 3, or a 4 into a 5 (possibly in multiple steps) and back again eventually. Notice that for $k = 5$, $n_1 = 19$. For example, say we begin with the edge $(1, 3, 4, 5, 6)$ and we want to switch a 3 in a particular spot into a 2, while maintaining the PSV composition. We will do this by first switching the 3 to a 5, then switching the 5 to a 2, in several steps:

- a. Switch any 2 into a 5 to create space between the number of 2's and the number of 3's, we will then have $\text{PSV}=(1, 2, 4, 5, 7)$;
- b. Switch the 3 that we want to eventually switch into a 2 into a 5, we then have $\text{PSV}=(1, 2, 3, 5, 8)$;
- c. Switch a *different* 5 back into a 3 to create space between the number of 2's and the number of 3's, $\text{PSV}=(1, 2, 4, 5, 7)$; and finally
- d. Switch the 5 we want into a 2, yielding the switch we originally wanted. $\text{PSV}=(1, 3, 4, 5, 6)$.

In general, our algorithm to reach the target sequence will then be to first, if we have more than one 1, change all extra 1's into k 's. We will then underline the single remaining 1 as something we won't touch again. Next, we will switch the number to the right of the 1 into a k , possibly in multiple steps, and then the k back into a 2, again possibly through multiple steps. We will now underline the 1 and the 2 together, as something we won't touch again. Next we will consider the next number to the right of this 2 and switch it to a k , and then switch back from the k into a 2 again, then underline the sequence 122, and switch all remaining 2's in the sequence into k 's. Next, we consider the number to right of the second 2, switch it into a k , and then switch from the k back into a 3, then block off 1223, etc., until we reach the target sequence. The following example illustrates the general technique.

Let $n = 19, k = 5$. Suppose we begin with the sequence 1432543552543455435 with $\text{PSV}=(1, 2, 4, 5, 7)$. There are no extra 1's to change into 5's. Then we can first change the 4 to the right of the 1 into a 5, but in order to do this, we must create space between the numbers of 3's and 4's by changing one 3 into a 5; we thus arrive at 1452543552543455435 $\text{PSV}=(1, 2, 3, 5, 8)$. The 4 to the right of the 1 can now be changed into a 5 to get 1552543552543455435 $\text{PSV}=(1, 2, 3, 4, 9)$. We now need to change the

5 to the right of the 1 back into a 2 by creating space between the number of 4's and 3's and then between the number of 3's and 2's. This leads to

$$1542543552543455435; (1, 2, 3, 5, 8);$$

$$1542343552543455435; (1, 2, 4, 5, 7);$$

and finally

$$\underline{12}42343552543455435; (1, 3, 4, 5, 6),$$

where the vectors above all represent PSV's. We then change the next 4 into a 5 and then back into a 2 as follows:

$$\underline{12}45343552543455435; (1, 2, 4, 5, 7),$$

$$\underline{12}45543552543455435; (1, 2, 3, 5, 8),$$

$$\underline{12}55543552543455435; (1, 2, 3, 4, 9),$$

$$\underline{12}54543552543455435; (1, 2, 3, 5, 8),$$

$$\underline{12}54343552543455435; (1, 2, 4, 5, 7),$$

and

$$\underline{122}4343552543455435; (1, 3, 4, 5, 6).$$

The rest of the algorithm proceeds as follows:

$$\underline{122}4343555543455435; (1, 2, 4, 5, 7),$$

$$\underline{122}4543555543455435; (1, 2, 3, 5, 8),$$

$$\underline{122}5543555543455435; (1, 2, 3, 4, 9),$$

$$\underline{122}5443555543455435; (1, 2, 3, 5, 8),$$

1223443555543455435; (1, 2, 4, 5, 7),

1223445555543455435; (1, 2, 3, 5, 8),

1223545555543455435; (1, 2, 3, 4, 9),

1223544555543455435; (1, 2, 3, 5, 8),

1223344555543455435; (1, 2, 4, 5, 7),

1223344555545455435; (1, 2, 3, 5, 8),

1223354555545455435; (1, 2, 3, 4, 9),

1223354455545455435; (1, 2, 3, 5, 8),

1223334455545455435; (1, 2, 4, 5, 7),

1223334455545455455; (1, 2, 3, 5, 8),

1223334445545455455; (1, 2, 3, 6, 7),

122333444555455455; (1, 2, 3, 5, 8),

1223334444555455455; (1, 2, 3, 6, 7),

122333444455555555; (1, 2, 3, 4, 9).

This process works in general since, given an edge of weight $n_1 := \frac{(k+4)(k-1)}{2} + 1$, the sums of the gaps between the components of the PSV may be as low as k , corresponding to the PSV $(1, 3, 4, \dots, k+1)$, or as high as $2k-1$, corresponding to the PSV $(1, 2, \dots, k-1, 2k-1)$. Switching numbers,

possibly in multiple steps, is always possible whenever there is a gap of two somewhere in the PSV sequence, which is guaranteed by the choice of n_1 . Question: Does a better algorithm allow for a smaller threshold n ? \square

The next theorem shows that there exists a uncovering of all partitions of $[n]$ into k parts if n is large enough; for simplicity we let the threshold n be the same as in Theorem 1. This is because any partition may be represented by a surjection satisfying the conditions of Theorem 2, though there may be multiple such representations when two or more of the part sizes are equal.

Theorem 2. *For each fixed $k, n; n \geq \frac{(k+4)(k-1)}{2} + 1$, there exists a ucycle of all onto functions $f : [n] \rightarrow [k]$ such that the preimage cardinality function $|f^{-1}| : \{\{1\}, \{2\}, \dots, \{k\}\} \rightarrow [n]$ is non-decreasing.*

Proof. Exactly the same as that of Theorem 1, except that the algorithm may terminate faster. \square

Corollary 3. *For any fixed $k \geq 3$, the unpacking and uncovering given in Theorems 1 and 2 respectively are both asymptotically of size $S(n, k)$, the number of partitions of $[n]$ into k parts.*

Proof. Our proof will reveal that

$$\frac{T(n, k)}{S(n, k)} = 1 - O\left(\frac{1}{\sqrt{n}}\right); \frac{U(n, k)}{S(n, k)} = 1 + O\left(\frac{1}{\sqrt{n}}\right), \quad n \rightarrow \infty,$$

where $T(n, k)$ and $U(n, k)$ are the lengths of the ucycles in Theorems 1 and 2 respectively. As pointed out by Professor László Székely, however, both these results are special cases of asymptotic results found in [5], where a threshold of $k = n^{1/5}$ is seen to hold for the property “partitions of size k with distinct parts form a “high” fraction of all partitions of $[n]$ into k parts.” Thus such values of k can serve to improve the conclusion of Corollary 3. Our proof is somewhat different, however.

Note that

$$\frac{T(n, k)}{S(n, k)} = 1 - \frac{Sa(n, k)}{S(n, k)},$$

where $Sa(n, k)$ denotes the number of partitions in which two or more parts are equal. Reframing the question in terms of distributing n distinguishable balls into k distinct boxes, we would like to calculate $\mathbb{P}(\cup_{i,j} I_{i,j})$ where $I_{i,j} =$

$\{B_i = B_j\}$, is the event that boxes i and j contain the same number of balls B_i and B_j respectively. By symmetry, we observe that

$$\begin{aligned} \mathbb{P}\left(\bigcup_{i,j} I_{i,j}\right) &\leq \binom{k}{2} P(B_1 = B_2) \\ &= \binom{k}{2} \sum_{j=0}^{\lfloor n/2 \rfloor} \frac{\binom{n}{j} \binom{n-j}{j}}{k^n} (k-2)^{n-2j} \\ &= \binom{k}{2} \left(\frac{k-2}{k}\right)^n \sum_{j=0}^{\lfloor n/2 \rfloor} \binom{n}{j} \binom{n-j}{j} \frac{1}{(k-2)^{2j}}. \end{aligned} \quad (1)$$

Lemma 4. *The function $f(j) = \binom{n}{j} \binom{n-j}{j} \frac{1}{(k-2)^{2j}}$ obtains its maximum value when $j = \frac{n}{k}$.*

Proof. Parametrize by setting $j = An$, we see (by Stirling's formula) that

$$\begin{aligned} &\binom{n}{An} \binom{(1-A)n}{An} \frac{1}{(k-2)^{2An}} \\ &= \frac{n!}{(An)!(An)!((1-2A)n)!} \frac{1}{(k-2)^{2An}} \\ &= \frac{K + o(1)}{n} \left(\frac{n}{e}\right)^n \left(\frac{e}{An}\right)^{2An} \left(\frac{e}{(1-2A)n}\right)^{(1-2A)n} \frac{1}{(k-2)^{2An}} \\ &= \frac{K}{nA^{2An}(1-2A)^{(1-2A)n}(k-2)^{2An}} (1 + o(1)). \end{aligned}$$

Maximizing this function is equivalent to minimizing the natural log of its denominator. Accordingly, define

$$\beta(A) = 2An \ln A + (1-2A)n \ln(1-2A) + 2An \ln(k-2).$$

Setting $\beta'(A) = 0$, we find easily that $A = \frac{1}{k}$. The next step is to show that $\frac{f(j+1)}{f(j)} < 1$ for $j > \frac{n}{k}$ and $\frac{f(j+1)}{f(j)} > 1$ for $j < \frac{n}{k}$. It is routine to calculate that

$$\frac{f(j+1)}{f(j)} = \frac{(n-2j)(n-2j-1)}{(k-2)^2(j+1)^2},$$

and thus that for any $\epsilon > 0$,

$$\frac{f\left(\frac{n}{k} + \epsilon + 1\right)}{f\left(\frac{n}{k} + \epsilon\right)} = \frac{\left(\frac{n}{k} - \frac{2\epsilon}{k-2}\right) \left(\frac{n}{k} - \frac{2\epsilon+1}{k-2}\right)}{\left(\frac{n}{k} + \epsilon + 1\right)^2} < 1,$$

and

$$\frac{f\left(\frac{n}{k} - \epsilon + 1\right)}{f\left(\frac{n}{k} - \epsilon\right)} = \frac{\left(\frac{n}{k} + \frac{2\epsilon}{k-2}\right)\left(\frac{n}{k} + \frac{2\epsilon-1}{k-2}\right)}{\left(\frac{n}{k} - \epsilon + 1\right)^2} > 1.$$

The lemma follows. \square

We now return to (1) and see that for a $\varphi(n)$ to be determined,

$$\begin{aligned} & \mathbb{P}\left(\bigcup_{i,j} I_{i,j}\right) \\ & \leq \binom{k}{2} \left(\frac{k-2}{k}\right)^n \left[\sum_{j=\frac{n}{k}-\varphi(n)}^{\frac{n}{k}+\varphi(n)} f\left(\frac{n}{k}\right) + \sum_{j=0}^{\frac{n}{k}-\varphi(n)} f(j) + \sum_{j=\frac{n}{k}+\varphi(n)}^{\lfloor n/2 \rfloor} f(j) \right] \\ & \leq \binom{k}{2} \left(\frac{k-2}{k}\right)^n \cdot \\ & \quad \left[2\varphi(n) f\left(\frac{n}{k}\right) + \left(\frac{n}{2} - 2\sqrt{n}\right) \max\left(f\left(\frac{n}{k} + \varphi(n)\right), f\left(\frac{n}{k} - \varphi(n)\right)\right) \right] (2) \end{aligned}$$

We will next use Stirling's approximation $N! \sim \sqrt{2\pi N}(N/e)^N$ at various points. Note that whenever $N = \frac{n}{k} + o(n)$ we have that $\sqrt{2\pi N} = \Theta\sqrt{n}$. Accordingly, we first see that for some constant A

$$\begin{aligned} f\left(\frac{n}{k}\right) & \leq \frac{A}{n} \left(\frac{n}{e}\right)^n \left(\frac{ke}{n}\right)^{2n/k} \left(\frac{ke}{(k-2)n}\right)^{(k-2)n/k} \left(\frac{1}{k-2}\right)^{2n/k} \\ & = \frac{A}{n} \left(\frac{k}{k-2}\right)^n, \end{aligned}$$

so that first component of (2) is no more than $B^{\frac{\varphi(n)}{n}}$. Next notice that

$$\begin{aligned}
& f\left(\frac{n}{k} + \varphi(n)\right) \\
& \leq \frac{C}{n} \left(\frac{n}{e}\right)^n \left(\frac{1}{k-2}\right)^{\frac{2n}{k} + 2\varphi(n)} \left(\frac{e}{n((1/k) + (\varphi(n)/n))}\right)^{\frac{2n}{k} + 2\varphi(n)} \\
& \quad \left(\frac{e}{n(1 - (2/k) - (2\varphi(n)/n))}\right)^{n - (2n/k) - 2\varphi(n)} \\
& = \frac{C}{n} \left(\frac{k-2}{k} + \frac{(k-2)\varphi(n)}{n}\right)^{-(2n/k) - 2\varphi(n)} \left(\frac{k-2}{k} - \frac{2\varphi(n)}{n}\right)^{-n + (2n/k) + 2\varphi(n)} \\
& = \frac{C}{n} \left(\frac{k}{k-2}\right)^n \left(1 + \frac{k\varphi(n)}{n}\right)^{-(2n/k) - 2\varphi(n)} \left(1 - \frac{2k}{k-2} \frac{\varphi(n)}{n}\right)^{-n + (2n/k) + 2\varphi(n)}.
\end{aligned}$$

Thus the second part of (2), when using $n/k + \varphi(n)$, is bounded above by

$$\begin{aligned}
& D \left(1 + \frac{k\varphi(n)}{n}\right)^{-(2n/k) - 2\varphi(n)} \left(1 - \frac{2k}{k-2} \frac{\varphi(n)}{n}\right)^{-n + (2n/k) + 2\varphi(n)} \\
& \leq E \exp \left\{ -2\varphi(n) - \frac{2k}{n} \varphi^2(n) + \frac{2k}{k-2} \varphi(n) - \frac{4}{k-2} \varphi(n) - \frac{4k}{k-2} \frac{\varphi^2(n)}{n} \right\} \\
& = E \exp \left\{ -\frac{2k}{n} \varphi^2(n) - \frac{4k}{k-2} \frac{\varphi^2(n)}{n} \right\},
\end{aligned}$$

which tends to zero provided that $\varphi(n) = \sqrt{n\psi(n)}$ for any $\psi(n) \rightarrow \infty$ (noting that $k \geq 3$ is fixed).

Finally it is easy to verify that the second part of (2) tends to zero if we consider $f(n/k - \varphi(n))$ as well. This completes the proof. \square

3 Universal Cycles of Partitions $\mathcal{P}(n, k)$ of $[n]$ into k parts

As in the previous section, we encode a k -partition of $[n]$ as a string of length n containing k symbols where i and j are in the same subset of the partition if and only if the i th character in the string is the same symbol as the j th character. Since the cases for $k = 1$ and $k = n$ are trivial, we always assume that $2 \leq k < n$. For convenience, we use $\{1, 2, \dots, k\}$ as our alphabet. We

refer to an encoding of a partition as a *representation* of that partition. Note that each k -partition of $[n]$ has $k!$ different representations.

Following methods outlined in [2], we construct a transition digraph $G_{n,k}$ for $\mathcal{P}(n, k)$ as follows. Let the set of vertices of $G_{n,k}$ be the set of all k and $(k-1)$ -partitions of $[n-1]$. There is an edge between two vertices v and w of $G_{n,k}$ if and only if w can immediately follow v in a ustring of k -partitions of $[n]$. That is, there is an edge from v to w if and only if the last $n-2$ symbols of a representation for v match the first $n-2$ symbols of a representation for w and the string formed by overlaying these two representations at their shared $n-2$ length substring is a representation of a k -partition of $[n]$. Observe that each vertex that is a k -partition of $[n-1]$ will have indegree = outdegree = k , and each vertex that is a $(k-1)$ -partition of $[n-1]$ will have indegree = outdegree = 1. As an example, $G_{5,3}$ is shown in Figure 1, with all vertices labeled with the representation having symbols appearing in the order 123.

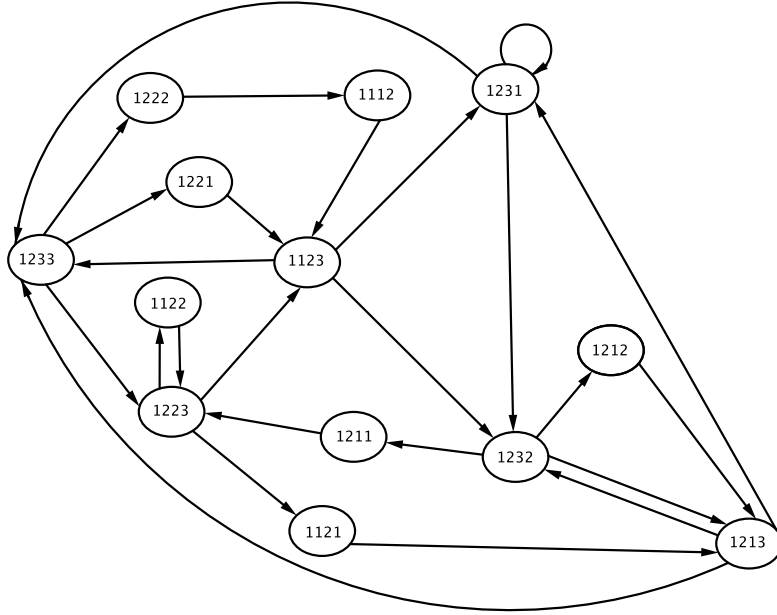


Figure 1: $G_{5,3}$

Now, the edges of $G_{n,k}$ are precisely the k -partitions of $[n]$. Furthermore, a partition p_1 can follow another partition p_2 in a ustring for $\mathcal{P}(n, k)$ if and only if the vertex at the tail of p_1 is also at the head of p_2 . Thus, there is a

bijection between the Eulerian cycles of $G_{n,k}$ and the ustrings of $\mathcal{P}(n, k)$.

Theorem 5. *Let $n, k \in \mathbf{Z}^+$ with $2 \leq k < n$, and let $G_{n,k}$ be the transition digraph for $\mathcal{P}(n, k)$. Then $G_{n,k}$ has an Eulerian cycle.*

Proof. $G_{n,k}$ is balanced as remarked above, so we must show that $G_{n,k}$ is weakly connected. To do so, we show that there exists a path from any vertex of $G_{n,k}$ to the vertex w with representation $(1, 2, \dots, k-1, k, k, \dots, k)$. Accordingly, let u be a vertex of $G_{n,k}$. We describe an algorithm for obtaining a path from u to w . We first find a path from u to a vertex v which ends in k distinct symbols. We may arrive at such a vertex in $k-1$ steps by a path $u = v_1, v_2, \dots, v_k$ where, for $i = 1, 2, \dots, k-1$, we choose v_{i+1} to be a vertex connected to v_i such that the representations of v_{i+1} end in $i+1$ distinct symbols. Note that choosing v_{i+1} this way is always possible - v_i will have representations ending in i distinct symbols and if $\text{outdegree}(v_i) = 1$ then the only possible choice for v_{i+1} has representations formed by adding the missing symbol of each representation of v_i to its last $n-2$ symbols (the case where $\text{outdegree}(v_i) = k$ is clear). Now, v_k has representations ending in k distinct symbols, so for any path of length $(n-1)-k$ starting at v_k , each vertex on the path will have $\text{outdegree} = k$. Thus, there exists a path $v_k, v_{k+1}, \dots, v_{n-1}$, where v_{k+j} has representations whose last $j+1$ symbols are all the same ($j = 0, 1, \dots, (n-1)-k$). Then, by construction, $v_{n-1} = w$. Hence, $G_{n,k}$ is weakly connected and so it follows that $G_{n,k}$ contains an Eulerian cycle. \square

Hence, we know that Eulerian cycles exist in $G_{n,k}$, and therefore ustrings of $\mathcal{P}(n, k)$ exist as well. However, there may be ustrings which cannot be turned into ucycles, which occurs when the representations of the first and last partitions do not overlap correctly, i.e., they have their symbols permuted. This idea is illustrated in the Eulerian cycle in Figure 2:

If we start with the representation 123 of the first vertex, then this Eulerian cycle represents the ustring 123312132, which cannot be turned into a ucycle. This example shows another important concept - once we choose the first representation to use, all other representations used are uniquely determined by the given Eulerian cycle. These observations motivate the following definitions.

Definition 6. *Suppose v is a vertex in $G_{n,k}$ and r is a representation for v . Form a new string r_0 from r by deleting all but the first occurrence of each symbol from r and appending the missing symbol to the end if v is a*

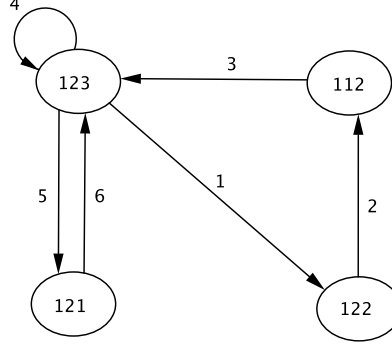


Figure 2: Eulerian cycle in $G_{4,3}$

$(k-1)$ -partition. Then r_0 is a permutation of $[k]$; call r_0 the relative order of r .

Definition 7. Consider an edge vw for some vertices v and w of $G_{n,k}$. Fix a representation r_v of v and suppose it has relative order π_v . Suppose the corresponding representation of w is r_w with relative order π_w . Then $\pi_w\pi_v^{-1}$ is called the associated permutation of the edge vw .

Remark 8. We have defined the associated permutation as the $\pi \in S_k$ such that $\pi\pi_v = \pi_w$, so that this definition is independent of the choice of representation of v .

The graph $G_{5,3}$ is shown again in Figure 3 with edges labeled with their associated permutations (expressed in cycle notation with fixed points suppressed).

Definition 9. Let $E = e_1, e_2, \dots, e_{S(n,k)}$ be an Eulerian cycle in $G_{n,k}$ and let π_i be the associated permutation of e_i , $i = 1, 2, \dots, S(n,k)$. We call the product $\pi_{S(n,k)}\pi_{S(n,k)-1} \cdots \pi_2\pi_1$ the permutation product of E .

From this definition, we get the following characterization.

Theorem 10. An Eulerian cycle $E = e_1, e_2, \dots, e_{S(n,k)}$ in $G_{n,k}$ can be lifted to a ucycle of $\mathcal{P}(n, k)$ if and only if its permutation product is the identity.

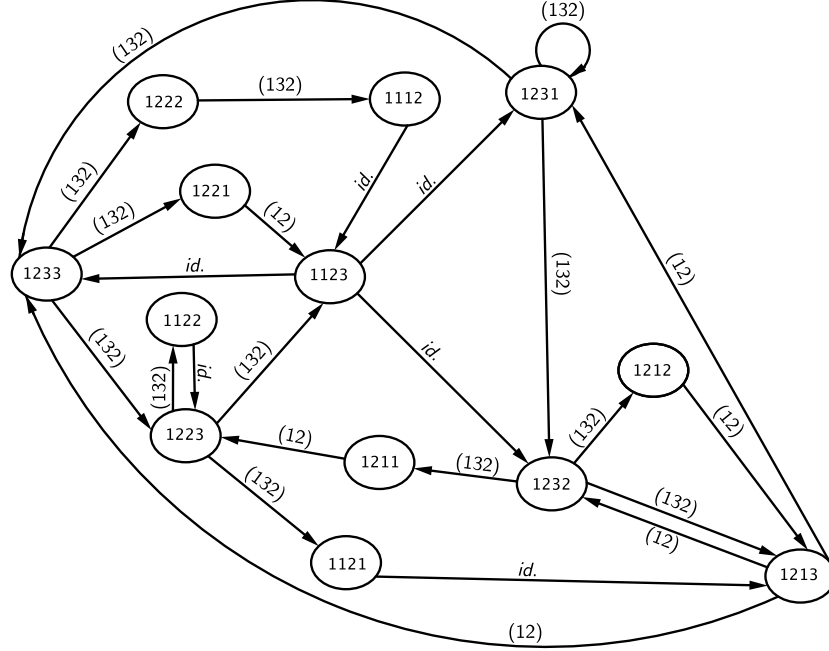


Figure 3: $G_{5,3}$ with associated permutations

Proof. Fix a representation r of the vertex at the tail of e_1 and suppose r has relative order τ . E can be lifted to a ucycle if and only if we arrive back at r at the end of the cycle, and going through E is equivalent to applying the permutation product to τ . \square

Now, we show that the associated permutation of an edge is completely determined by the vertex at its “tail”, and that only certain permutations can be associated permutations.

Lemma 11. *Let vw_1 be an edge in $G_{n,k}$, suppose vw_1 has associated permutation π . Then π has the form $(1\ j\ j-1 \cdots 2)$ for some $1 \leq j \leq k$, and if vw_2 is another edge from v , then vw_2 has associated permutation π as well.*

Proof. Let $r = v_1v_2 \cdots v_{n-1}$ be the representation of v with relative order $12 \cdots k$. Then there are representations $r_1 = v_2 \cdots v_{n-1}u_1$ and $r_2 = v_2 \cdots v_{n-1}u_2$ of w_1 and w_2 corresponding to r under vw_1 and vw_2 , respectively. Suppose that $j-1$ distinct symbols appear in r after v_1 ($= 1$) and before a second appearance of 1 (1 may only occur once in r). Since the first $n-2$

characters of both r_1 and r_2 are the same as the last $n - 2$ characters of r , it follows that r_1 and r_2 both have relative order $23 \cdots j \ 1 \ j+1 \ j+2 \cdots k$. Hence, the associated permutations of vw_1 and vw_2 are both $(1 \ j \ j-1 \cdots 2)$. \square

Theorem 12. *For $n \geq 3$, every Eulerian cycle of $G_{n,2}$ can be lifted to a ucycle.*

Proof. Observe that the vertex with representation $11 \cdots 1$ is the only vertex of outdegree 1 and that the edge coming out of this vertex has $id.$ as its associated permutation. All other vertices have outdegree 2, and the two edges originating from any particular vertex both have the same associated permutation by Lemma 11. In particular, there is an even number of (12) permutations. Since S_2 is abelian, the permutation product of an Eulerian cycle will be the identity so that the result follows by Theorem 10. \square

Corollary 13. *For $n \geq 3$, ucycles of $\mathcal{P}(n, 2)$ exist.*

Proof. This follows directly from Theorem 5 and Theorem 12. \square

We can also use the permutation product to determine cases when ucycles do not exist. The easiest way for this to occur is if the multiset consisting of all associated permutations in $G_{n,k}$ contains an odd number of odd permutations since this ensures that there is no ordering of the associated permutations which multiplies to the identity.

Definition 14. *We call the multiset consisting of all associated permutations in $G_{n,k}$ the permutation multiset of $G_{n,k}$.*

Definition 15. *Let \mathcal{O} be the multiset which contains all odd permutations of the permutation multiset of $G_{n,k}$. Define the parity function by*

$$Par(n, k) = \begin{cases} 0 & \text{if } |\mathcal{O}| \equiv 0 \pmod{2} \\ 1 & \text{if } |\mathcal{O}| \equiv 1 \pmod{2} \end{cases}$$

Lemma 16. *If $Par(n, k) = 1$, then there does not exist a ucycle of $P(n, k)$.*

The following formula gives a recursive formula for calculating $Par(n, k)$.

Lemma 17. *The function $Par(n, k)$ satisfies the following recurrence relation:*

$$Par(n, k) \equiv k \cdot Par(n-1, k) + Par(n-1, k-1) + S(n-2, k-2) \pmod{2} \quad (3)$$

with initial conditions $Par(n, 2) = 0$ for all n , and

$$Par(n, n-1) = \begin{cases} 1 & \text{if } n \equiv 0 \pmod{4} \\ 0 & \text{otherwise} \end{cases}$$

Proof. We establish a relationship between the permutation multiset of $G_{n,k}$ and those of $G_{n-1,k}$ and $G_{n-1,k-1}$. Suppose v is a vertex in $G_{n,k}$, so v represents a k or $k-1$ -partition p of $[n-1]$. We consider the edge e_p which represents p in either $G_{n-1,k}$ or $G_{n-1,k-1}$. We know that the associated permutation of e_p is determined by the location of the second occurrence of the first symbol in a representation r_p of the vertex w_p at the tail of e_p by Lemma 11. First, suppose v represents a k -partition of $[n-1]$. If the first symbol does actually occur for a second time in r_p , then since there is a representation of v whose first $n-2$ characters are precisely r_p , it follows that e_p has the same associated permutation as all the edges coming from v . If the first symbol of r_p does not occur a second time, then e_p has associated permutation $(1 \ k \ k-1 \cdots 2)$. If w_p has outdegree 1, then the representations of v do not have a second occurrence of their first symbols, and so all edges from v have associated permutation $(1 \ k \ k-1 \cdots 2)$. If w_p has outdegree k , then the representations of v have an occurrence of all symbols before a second occurrence of the first symbol, so we get that the edges from v have associated permutation $(1 \ k \ k-1 \cdots 2)$ again. Since each vertex of $G_{n,k}$ which represents a k -partition of $[n-1]$ has outdegree k , we get the term $k \cdot Par(n-1, k)$.

Now, suppose v represents a $k-1$ -partition of $[n-1]$.

Case 1, the first symbol of v appears a second time. Then either the first symbol of r_p appears a second time, or the first symbol of r_p is appended by following e_p (if the second appearance in v is at the last character). If the first symbol of r_p appears a second time, then by previous reasoning e_p has the same associated permutation as all edges from v . If the first symbol of r_p does not occur a second time, then e_p and the edges from v all have associated permutation $(1 \ k \ k-1 \cdots 2)$.

Case 2, the first symbol of v does not appear a second time. Then the first symbol of r_p does not appear a second time, and so e_p must have associated permutation $(1 \ k-1 \ k-2 \cdots 2)$. However, in this case v has associated permutation $(1 \ k \ k-1 \cdots 2)$. Note that since the first symbol of v does not appear a second time, the last $n-2$ characters of v represent a $k-2$ -partition of $[n-2]$, so this case occurs exactly $S(n-2, k-2)$ times. Thus, we have $S(n-2, k-2)$ partitions that either switch from even to odd or odd to even;

in either case adding $S(n-2, k-2)$ affects the parity in the desired manner.

Thus, each vertex in $G_{n,k}$ which represents a $k-1$ -partition of $[n-1]$ has the same associated permutation as it does in the graph $G_{n-1,k-1}$ except for $S(n-2, k-2)$ permutations which change sign. Since each such vertex has outdegree 1, we get the $Par(n-1, k-1) + S(n-2, k-2)$ term.

Finally, the initial condition $Par(n, 2) = 0$ for all n follows from Theorem 12, and the initial condition

$$Par(n, n-1) = \begin{cases} 1 & \text{if } n \equiv 0 \pmod{4} \\ 0 & \text{otherwise} \end{cases}$$

follows from the remark following Theorem 5 in [4] □

Corollary 18. *For $n \geq 4$ and $2 \leq k < n$,*

$$Par(n, k) \equiv \begin{cases} 0 \pmod{2} & \text{if } n \text{ is odd} \\ S(n-2, k-2) \pmod{2} & \text{if } n \text{ is even} \end{cases} \quad (4)$$

Proof. We proceed by induction. For the base cases, we show that the initial conditions of the recursion in Corollary 17 satisfy Equation (4). If we define $S(n, 0) := 0$ for all n , then Equation (4) yields $Par(n, 2) = 0$ for all n . Also, if n is even, then Equation 4 yields

$$\begin{aligned} Par(n, n-1) = S(n-2, n-3) &= \binom{n-2}{2} = \frac{(n-2)(n-3)}{2} \\ &\equiv \begin{cases} 0 \pmod{2} & \text{if } n \equiv 2 \pmod{4} \\ 1 \pmod{2} & \text{if } n \equiv 0 \pmod{4} \end{cases} \end{aligned}$$

as desired.

Now, suppose that (4) holds for $Par(n-1, k)$ and $Par(n-1, k-1)$. If n is odd then $n-1$ is even, so by (3), the induction hypothesis, and the fundamental Stirling number recurrence,

$$\begin{aligned} Par(n, k) &\equiv k \cdot Par(n-1, k) + Par(n-1, k-1) + S(n-2, k-2) \pmod{2} \\ &\equiv k \cdot S(n-3, k-2) + S(n-3, k-3) + S(n-2, k-2) \pmod{2} \\ &\equiv 2S(n-3, k-2) + S(n-2, k-2) + S(n-2, k-2) \pmod{2} \\ &\equiv 2S(n-2, k-2) \pmod{2} \\ &\equiv 0 \pmod{2} \end{aligned}$$

If n is even, then $n - 1$ is odd and so $Par(n - 1, k) = Par(n - 1, k - 1) = 0$. Hence, $Par(n, k) \equiv S(n - 2, k - 2) \pmod{2}$ by (3). \square

Corollary 19. *If n is even and $S(n - 2, k - 2)$ is odd, there does not exist a ucycle for k -partitions of $[n]$.*

Proof. This follows directly from Corollary 18. \square

Corollary 20. *If $n \geq 4$ is even, then there does not exist a ucycle for 3, 4, or 6-partitions of n .*

Proof. By Corollary 19, we must prove that $S(n, 1)$, $S(n, 2)$, and $S(n, 4)$ are odd for all even n . We know $S(n, 1) = 1$ and $S(n, 2) = 2^{n-1} - 1$, which are both odd. Now,

$$\begin{aligned} S(n, 4) &= \frac{1}{4!} \left[-\binom{4}{1}1^n + \binom{4}{2}2^n - \binom{4}{3}3^n + \binom{4}{4}4^n \right] \\ &= \frac{1}{4!}(-4 + 6 \cdot 2^n - 4 \cdot 3^n + 4^n) \\ &= \frac{1}{6}(-1 + 3 \cdot 2^{n-1} - 3^n + 4^{n-1}) \end{aligned}$$

If we consider the term in parentheses mod 4, then we get $(-1 + 0 - 1 + 0) \equiv 2$ ($3^n \equiv 1 \pmod{4}$ for even n). Thus 2 only divides the term in parentheses once, so when we divide it by 6 we get an odd number. \square

4 Open Questions

In addition to the question raised in the last line of the proof of Theorem 1, we can ask the following:

(a) The smallest case remaining after our investigation is the one for $n = 5, k = 3$, for which we have found the ucycle 3112311123213233112131322. All other cases for $n = 5$ and $n = 6$ are solved. This leads to the question: What is the best result that can be proved along the following lines “For $n \geq 3$ and $3 \leq k < n$, ucycles of k -partitions of $[n]$ exist if and only if n is odd?”

(b) In general, how can one use Theorem 10 to prove results on *existence* of ucycles (rather than non-existence)?

- (c) Even if ucycles of $\mathcal{P}(n, k)$ may not exist, when is it true that ucycles exist for $\mathcal{P}(n, s, t)$, the set of partitions of $[n]$ into between s and t parts; $s < t$?
- (d) Throughout this paper, we have insisted on having the alphabet size equal k . How do our results change if we relax this condition?

5 Acknowledgments

The research of all the authors was supported by NSF Grant 1263009.

References

- [1] A. Blanca and A. Godbole (2011). “On universal cycles for new classes of combinatorial structures,” *SIAM J. Discrete Math.* **25**, 1832–1842.
- [2] F. Chung, P. Diaconis, and R. Graham (1992). “Universal cycles for combinatorial structures,” *Discrete Math.* **110**, 43–59.
- [3] D. Curtis, T. Hines, G. Hurlbert, and T. Moyer (2009). “Near-universal cycles for subsets exist,” *SIAM J. Discrete Math.* **23**, 1441–1449.
- [4] K. Casteels and B. Stevens (2009). “Universal cycles of $(n-1)$ -partitions of an n -set,” *Discrete Math.* **309**, 5332–5340.
- [5] É. Czabarka, M. Marsili, and L. A. Székely (2013). “Threshold functions for distinct parts: revisiting Erdős–Lehner,” in: *Information Theory, Combinatorics, and Search Theory (in Memory of Rudolph Ahlswede)*, eds. H. Aydinian, F. Cicalese, C. Deppe, Lecture Notes in Computer Science **7777**, Springer-Verlag, 463–471.
- [6] M. Dewar and B. Stevens (2102). *Gray Codes, Universal Cycles, and Configuration Orderings*, Springer Verlag, New York.
- [7] A. Elks, A. Godbole, and S. McInturff (2014+). “Some results on universal cycles for partitions of a set,” Preprint.
- [8] G. Hurlbert (1994). “On universal cycles for k -subsets of an n -set,” *SIAM J. Discrete Math.* **7**, 598–604.

- [9] B. LaBounty-Lay, A. Bechel, and A. Godbole (2008). “Universal cycles of discrete functions,” *Cong. Numer.* **189**, 121–128.